

AN ENHANCED FACIAL EXPRESSION CLASSIFICATION SYSTEM USING EMOTIONAL BACK PROPAGATION ARTIFICIAL NEURAL NETWORK WITH DCT APPROACH

VEDANTHAM RAMACHANDRAN¹, E. SRINIVASA REDDY² & B. SIVIAIAH³

¹Research Scholar, Department of CSE, Acharyanagarjuna University, Nambur, Guntur, Andhra Pradesh, India

²Professor & Principal, Acharya Nagarjuna University, Nambur, Guntur, Andhra Pradesh, India

³Assistant Professor, Vignan's University, Vadlamudi, Guntur, Andhra Pradesh, India

ABSTRACT

In this paper a novel classification system for facial expressions with back propagation artificial neural network that uses Discrete Cosine Transformation technique for pre-processing in feature extraction of images is proposed. This technique is capable of processing images extremely rapidly while achieving high detection rates for facial expressions. This entire system deploys two important contributions that are optimally mixed to make the entire system swifter. The first is the selection of an image representation that incorporates DCT for the relatively quick feature extraction suitable for the facial databases. The second is a scrupulously modified Emotional Back Propagation Neural Network classifier to select a subset of typical decisive visual features from a set of potential features. The optimum mixture of the above two techniques facilitated quick learning and there by computational speeds deducing the promising facial expressions. A set of experiments in the domain of facial expression detection were conducted. The system yields facial expression classification performance comparable to the previous systems that are implemented on a standard desktop.

KEYWORDS: Facial Emotion Recognition, DCT, Back Propagation Neural Network

INTRODUCTION

Facial expressions, and other gestures, convey non-verbal communication cues in face-to-face interactions. Recently, the facial expression recognition technology attracts more and more attention with people's growing interest in expression information. Facial expression carries crucial information about the mental, emotional and even physical states of the conversation. Facial expression recognition has practical significance; it has very broad application prospects, such as user-friendly interface between man and machine, humanistic design of goods, and emotional robot etc. With facial expression recognition systems, the computer will be able to assess the human expressions depending on their effective state in the same way that human's senses do. The intelligent computers will be able to understand, interpret and respond to human intentions, emotions and moods [4].

Computer recognition of human face identity is the most fundamental problem in the field of pattern analysis. Emotion analysis in man-machine interaction system is designed to detect human face in an image and analyse the facial emotion or expression of the face.

This helps in improving the interaction between the human and the machine. The machines can thereby understand the man's reaction and act accordingly. This reduces the human work hours. For example, robots can be used as a class tutor, pet robots, CU animators and so on.. We identify facial expressions not only to express our emotions, but also to provide important communicative cues during social interaction, such as our level of interest, our desire to take a speaking turn and continuous feedback signalling or understanding of the information conveyed. Support Vector Algorithm is well suited for this task as high dimensionality does not affect the Gabor Representations. The main disadvantage of the system is that it is very expensive to implement and maintain. Any changes to be upgraded in the system needs a change in the algorithm which is very sensitive and difficult; hence our developed system will be the best solution to overcome the abovementioned disadvantages [5].

In this paper, we describe a classification system for facial expressions with back propagation artificial neural network that uses Discrete Cosine Transformation technique for pre-processing in feature extraction of images. This technique is capable of processing images extremely rapidly while achieving high detection rates for facial expressions. This entire system deploys two important contributions that are optimally mixed to make the entire system swifter.

The first is the selection of an image representation that incorporates DCT for the relatively quick feature extraction suitable for the facial databases. The second is a scrupulously modified Emotional Back Propagation Neural Network classifier to select a subset of typical decisive visual features from a set of potential features. Our proposed system The system automatically detects and extracts the human face from the background based on a combination of a retainable neural network structure. In this system, the computer is trained with the various emotions of the face and when given an input, the computer detects the emotion by comparing the co-ordinates of the expression with that of the training examples and produces the output. Principle Component Analysis algorithm is the one being used in this system to detect various emotions based on the coordinates of the training sample given to the system.

The Following figure shows a model for Face expression Detection system which uses the Image data base and compares the input image with the images in the database by several algorithms and then it recognises the image.

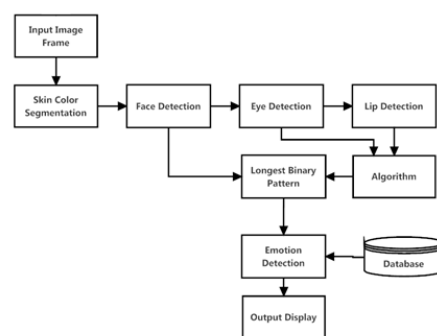


Figure 1: An Example of Image Recognition System

PROPOSED FACIAL RECOGNITION SYSTEM

This work proposes a new solution to the facial expression recognition problem, describing a facial recognition system that can be used in application of Human computer interface. The proposed architecture in this work contains the following stages: reprocessing of input images, feature extraction, training, classification, and database.

Pre-processing of input images includes, face detection and cropping. Feature extraction is the process of deriving unique features from the data and can be accomplished by specific algorithms like Feature averaging, principal component analysis etc. Training of neural network will be done by giving the extracted features as input to the neural network with specified network parameters. Classification will be done by the neural network according to the specified targets in the network.

The following figure shows the proposed architecture of the Facial recognition system. Pre-processing is the next stage after entering the data into the facial expression recognition system. The important data that is needed for most facial expression recognition methods is face position. In pre-processing module images are resized from 256 x 256 pixel value to 280 x 180 pixel values. The Sobel method has been used to identify the face edges.

The proposed architecture in this work contains the following stages: preprocessing of input images, feature extraction, training, classification, and database. Preprocessing of input images includes, face detection and cropping. Feature extraction is the process of deriving unique features from the data and can be accomplished by specific algorithms like Feature averaging, principal component analysis etc. Training of neural network will be done by giving the extracted features as input to the neural network with specified network parameters. Classification will be done by the neural network according to the specified targets in the network.

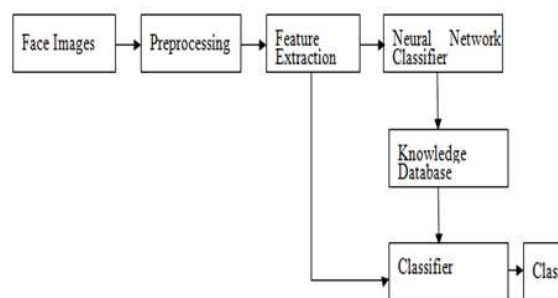


Figure 2: Architecture of the Proposed Method

Face images are taken from Cohn Kanade database of facial expressions. The original image contains time and camera model also. For better performance, face is detected and cropped and saved as separate image. The cropped image is then used to extract features. These features are given as input to the neural network and will be trained to gain knowledge.

Preprocessing

The testing image will also be preprocessed and features will be extracted and input to the neural network. The classifier of the neural network will classify the expression of the input test image.

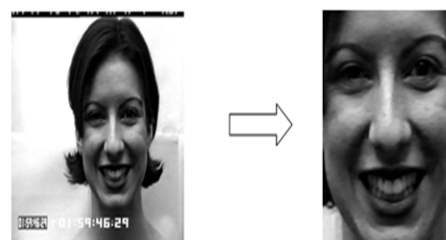


Figure 3: Face Detection and Cropping

In order to perform data reduction, the first step is to take the required data from an image. So the face is detected and cropped from original image as shown in Figure 3.

Discrete Cosine Transform

DCT is the most widely used transform in the image processing applications for feature extraction. The approach involves taking the transformation of the image as a whole and separating the relevant coefficients. DCT performs energy compaction [1]. The DCT of an image basically consists of three frequency components namely low, middle, high each containing some detail and information in an image. The low frequency generally contains the average intensity of an image which is the most intended in FR systems [2]. Mathematically, the 2D-DCT of an image is given by:

$$F(p, q) = \alpha(p) \alpha(q) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \cos\left[\frac{\pi}{2N}(2x+1)\right] \cos\left[\frac{\pi}{2M}(2y+1)\right] f(x, y) \quad (1)$$

$$\alpha(p) \alpha(q) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } p, q \neq 0 \\ \sqrt{\frac{2}{N}} & \text{for } p, q = 0 \end{cases} \quad (2)$$

where $f(x, y)$ is the intensity of the pixel at coordinates (x, y) , u varies from 0 to $M-1$, and v varies from 0 to $N-1$, where $M \times N$ is the size of the image.

DCT Coefficients with Feed Forward Neural Network

This architecture specifies the classification with neural network using pattern averaging of input images. The training images were taken and applied the pattern averaging. The remaining features are input to the feed forward neural network to train the network.

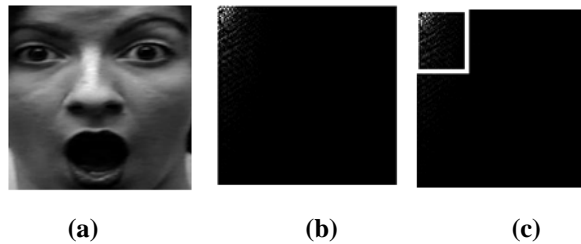


Figure: 4 (a) Image from Cohn Canede Expression Database (b) Its DCT Transformed Image (c)Top-Left (Low Frequency) Rectangle Carries Maximum Information

The neural network will produce the knowledge database. In the process of testing, the test input image will be applied pattern averaging and the remaining features will be used to classify through the neural network classifier and using the knowledge database gained from training. The architecture is shown in the Figure 5.

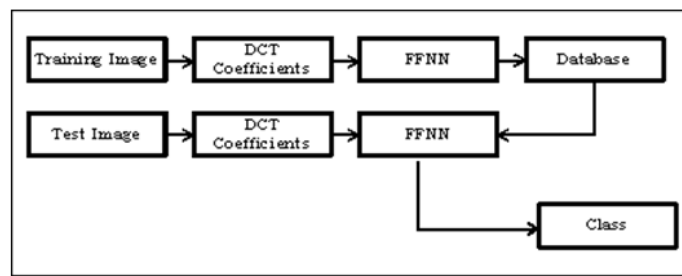


Figure 5: Architecture of FFNN Classification with DCT Coefficients

DCT Coefficients with Emotional Back Propagation Neural Network

This architecture proposes the classification with neural network using pattern averaging of input images. The training images were taken and applied the pattern averaging. The remaining features are input to the emotional back propagation neural network to train the network. The neural network will produce the knowledge database.

In the process of testing, the test input image will be applied pattern averaging and the remaining features will be used to classify through the neural network classifier and using the knowledge database gained from training. The architecture is shown in the Figure 6.

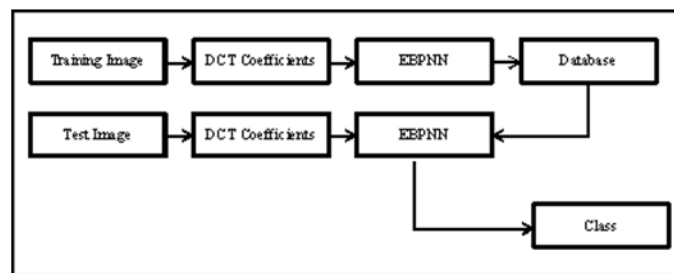


Figure 6: Proposed Architecture of EBPNN Classification with DCT Coefficients

The proposed two architectures uses the emotional back propagation neural network architecture. The generalized architecture of the proposed system is shown in Figure 7.

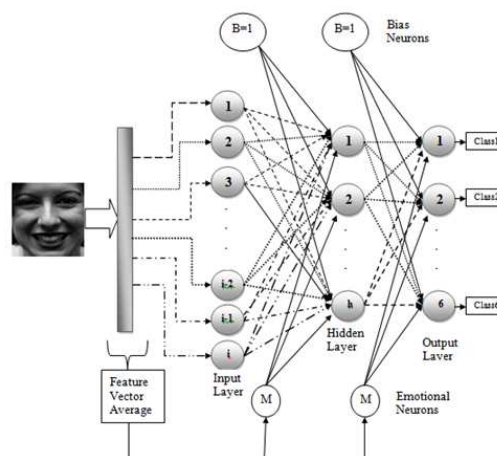


Figure 7: Generalized EMBP-Based Neural Network

TESTS, RESULTS & CONCLUSIONS

The implementation of neural network consists of training and testing. The training and testing is performed on Cohn Kanade facial expression database. The database consists of 2000 images of 200 subjects. About 600 images were used in this work for the training and testing process. Sample images from the Cohn Kanade database are shown in Figure 4.

The performance of the system is measured by varying the number of images of each expression in training and testing. Following table shows the performance of the proposed method along with the other methods.

Table 1: Comparison of Results on Cohn Kanade Database

Training Samples	Testing Samples	Training Time		Recognition Rate	
		DCT+FFNN	DCT+BPNN	DCT+FFNN	DCT+BPNN
8	2	160.85	152.30	100	100
7	3	144.29	202.39	100	100
6	4	69.50	84.43	96	98
5	5	25.145	33.69	87	97
4	6	8.86	13.08	85	96
3	7	3.37	4.48	83	95
2	8	2.88	4.17	80	95

The recognition performance increases as the number of training samples increases. The lower the number of training samples the lesser the recognition rate. It is found that the DCT with emotional back propagation neural network is yielding the better results even the training samples are less. The performance plot was shown against various algorithms, number of training images and their performances.

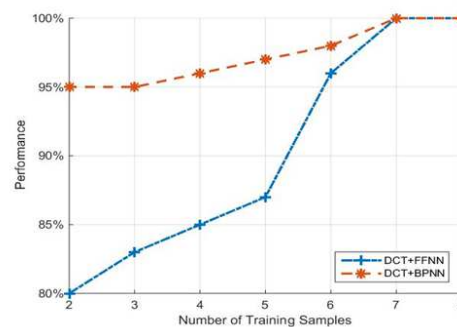


Figure 8: Plot against FFNN & BPNN Recognition Rate

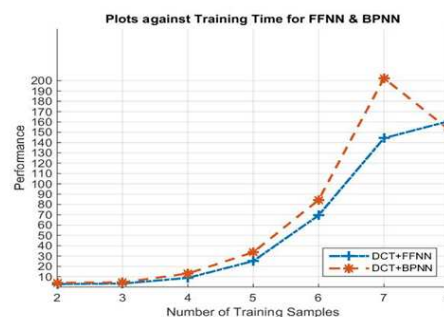


Figure 9: Plot against FFNN & BPNN Training Times

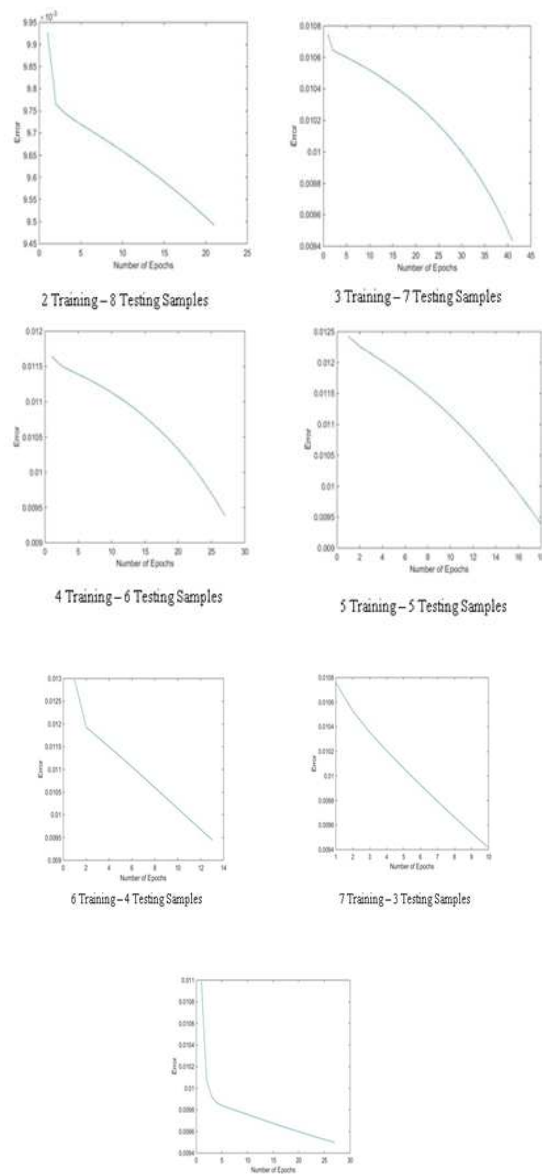


Figure 10: Error Minimization Plots for EBPNN

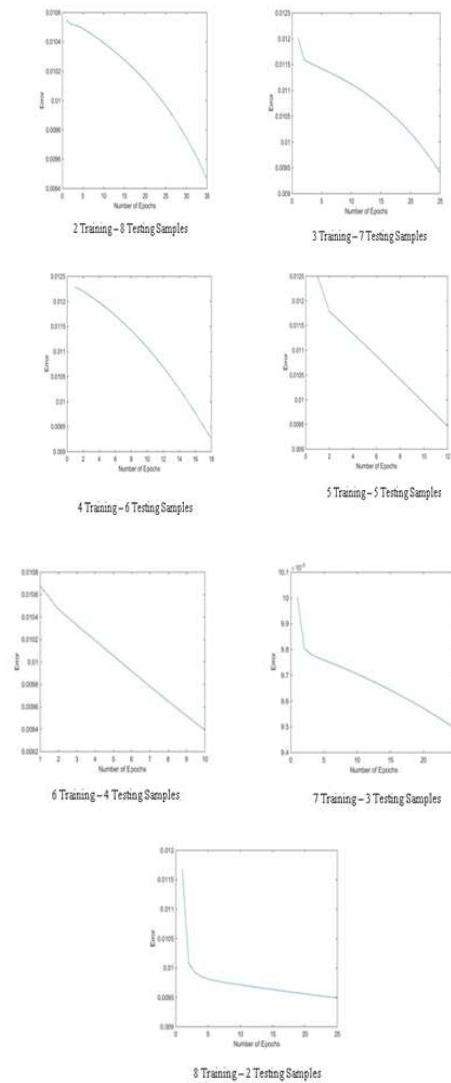


Figure 11: Error Minimization Plots for Ffn

The confusion matrix is created for each of the test. The test is performed on five subjects.

Found	Surprise	Fear	Happy	Sad	Anger	Disgust
Actual						
Surprise	100	0	0	0	0	0
Fear	0	100	0	0	0	0
Happy	0	0	60	0	0	40
Sad	0	0	0	100	0	0
Anger	0	0	0	60	40	0
Disgust	0	0	0	0	0	100

Figure 12: DCT+BPNN Confusion Matrix

The confusion matrix shows the percentage of correct classifications and mis-classifications also. Diagonal elements show the correct classification results.

All other elements are the misclassifications. The test and training results of various facial emotional classification methods is shown in Figure. Experimental results show that the proposed architecture improves the performance of the

facial expressions. Based on the results we can conclude that the proposed emotional back propagation neural network with DCT is best in both cases of minimization of training time of neural network and performance as well. Since the emotional parameters were introduced, the training time for the single iteration may be little more but the overall training time is reduced in achieving the minimization of error.

The performance and training time of the neural network depends on the parameters selected like learning coefficient and momentum factor. The number of hidden neurons is also affecting the performance of the neural network. Experiments were carried out by altering the learning coefficient and number of hidden neurons and the types of sigmoid functions.

The optimal value for learning rate is 0.02, which produces the best performance for facial expression recognition. The number of hidden neurons is same as the number of input neurons. Sigmoid action function is used in both hidden layer and output layer for activating the neurons. In the classification part of the emotional back propagation neural network, the time very less when compared to other neural networks.

The work can be extended to clustering techniques like segmentation for the lower training times and higher performance. Since the training data is still images, there is more dependency on the image data like lighting, illumination conditions, poses of the faces, variations in expression and gender of the person also.

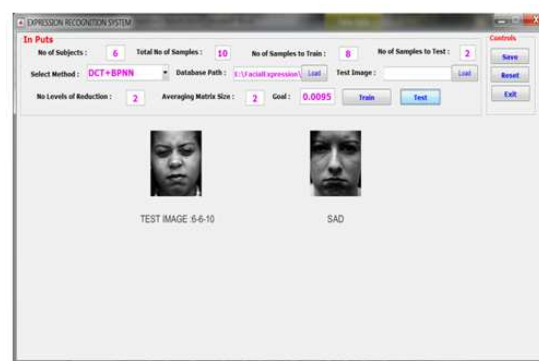


Figure 13: GUI Result for the Developed System

CONCLUSIONS AND FUTURE WORK

We present an approach for facial expression estimation that combines state-of-the-art techniques for model-based image interpretation and sequence labelling. Learned objective functions ensure robust model-fitting to extract accurate model parameters. The classification algorithm is explicitly designed to consider sequences of data and therefore considers the temporal dynamics of facial expressions. Future work aims at presenting the classifier training data that is obtained from various publicly available databases to reflect a broader. Variety of facial expressions. Furthermore, our approach will be tuned towards applicability in real-time. It is planned to create a working demonstrator from this approach.

REFERENCES

1. Adnan Khashman , “A modified back propagation learning algorithm with added emotional coefficients”, IEEE transactions on neural networks, vol. 19, no. 11, November 2008
2. D. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning internal representations by error propagation,” in Parallel Distributed Processing, D. Rumelhart and J. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1.

3. H. Murase and S. K. Nayar, "Visual learning and recognition of 3-D objects from appearance," *Int. J. Comput. Vis.*, vol. 14, pp. 5–24, 1995.
4. P. M. Chavan, Manan C .Jadhav, Jinal B. Mashruwala, Aditi K. Nehete, Pooja A. Panjari, "Real Time Emotion Recognition through Facial Expressions for Desktop Devices", *International Journal of Emerging Science and Engineering (IJESE)* Volume-1, Issue-7, May 2013
5. Ajit P. Gosavi, S. R. Khot, "Facial Expression Recognition Using Principal Component Analysis", *International Journal of Soft Computing and Engineering (IJSCE)*, Volume-3, Issue-4, September 2013
6. V.Ramachandran, et.al "Facial Expression Classification Systems with Emotional Back Propagation Neural Network " *International Journal of Scientific and Engineering Research (IJSER)* - (ISSN 2229-5518), Volume 4, Issue 9.during Aug 2013
7. Lt. Dr. S. Santhosh Baboo, Lt. Dr. S. Santhosh Baboo," Face Emotion Analysis Using Gabor Features In Image Database for Crime Investigation", *International Journal of Data Engineering (IJDE)*, Volume (2) : Issue (2) : 2011
8. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
9. S. Z. Li, X. W. Hou, H. J. Zhang, and Q. S. Cheng, "Learning Spatially localized, parts-based representation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. 207–212.
10. L. L. Huang and A. Shimizu, "Combining classifiers for robust face detection," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2006, vol. 3972, pp. 116–121.
11. C. Park, M. Ki, J. Namkung, and J. K. Paik, "Multimodal priority verification of face and speech using momentum back-propagation neural network," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2006, vol. 3972, pp. 140–149.

AUTHORS DETAILS



Vedantham Ramachandran, is a research scholar at Acharya Nagarjuna University, Nambur and an associate professor at Vasireddy Venkatadri Institute of Technology. He did B.Tech Computer Science & Systems engineering Degree from Andhra University and M.Tech in Computer Science Engineering from JNTU, Kakinada. He is very much interested in image processing, medical image retrieval, human vision & pattern recognition. He did several projects in image processing.



Dr. E. Srinivasa Reddy, Professor & Dean R&D in University Engineering College of Acharya Nagarjuna University. He did Master of Technology in Computer Science Engineering from sir Mokhsagundam Visweswariah University, Bangalore. He also did Mater of Science from Birla Institute of Technological Sciences, Pilani. He did Philosophical Doctorate in Computer Science Engineering from Acharya Nagarjuna University, Guntur. Prof.E.S.Reddy, guided several research scholars of PhD to success. His research areas of interest are in various domains including Digital Image Processing.



B. Sivaiah received the B. Tech. degree in Computer Science & Engineering from Jawaharlal Technological University, Hyderabad, Andhra Pradesh, India in 2007. He got M.Tech. degree in Computer Science & Engineering from Jawaharlal Technological University, Kakinada, Andhra Pradesh, India in 2010. Presently he is working as assistant professor in Department of Computer Science & Engineering, Vignan's University, Andhra Pradesh, India. His current research interests include image processing, neural networks, machine learning, intelligent systems, pattern recognition and artificial intelligence.

